

Comparing the Treatment of Uncertainty in Bayesian Networks and Fuzzy Expert Systems Used for a Human Reliability Analysis Application

Piero Baraldi^{a*}, Luca Podofillini^b, Lusine Mkrtchyan^b, Enrico Zio^{a,c}, Vinh N. Dang^b

^aEnergy Department, Polytechnic of Milan, Milan, Italy

^bPaul Scherrer Institute, PSI Villigen, Switzerland

^cChair on Systems Science and the Energetic Challenge, European Foundation for New Energy -
Électricité de France, École Centrale Paris and Supélec

ABSTRACT: The use of expert systems can be helpful to improve the transparency and repeatability of assessments in areas of risk analysis with limited data available. In this field, Human Reliability Analysis (HRA) is no exception, and, in particular, dependence analysis is an HRA task strongly based on analyst judgement. The analysis of dependence among Human Failure Events refers to the assessment of the effect of an earlier human failure on the probability of the subsequent ones. This paper analyses and compares two expert systems, based on Bayesian Belief Networks and Fuzzy Logic (a Fuzzy Expert System, FES), respectively. The comparison shows that a BBN approach should be preferred in all the cases characterized by quantifiable uncertainty in the input (i.e. when probability distributions can be assigned to describe the input parameters uncertainty), since it provides a satisfactory representation of the uncertainty and its output is directly interpretable for use within PSA. On the other hand, in cases characterized by very limited knowledge, an analyst may feel constrained by the probabilistic framework, which requires assigning probability distributions for describing uncertainty. In these cases, the FES seems to lead to a more transparent representation of the input and output uncertainty.

Keywords: expert judgement; expert models; Bayesian belief networks; fuzzy fogic; human reliability analysis; dependence assessment;

* piero.baraldi@polimi.it; Tel: +390223996355; Fax: +390223996309

1 Introduction

Handling of different sorts of uncertainties is intrinsic to risk analysis and, correspondingly, many publications have addressed the issue since decades [1-4]. The probabilistic and Bayesian frameworks are since long the established practice to represent and quantify uncertainty in risk analysis and, in particular, in Probabilistic Safety Assessment (PSA) [5-7]. Recently, alternative, non-probabilistic frameworks (e.g. Fuzzy Logic, possibility theory, imprecise probability) have also been increasingly proposed and investigated to represent uncertainty in situations characterized by limited available information, mainly coming from expert judgements [8-13]. However, these frameworks have not yet found visible application outside the scientific community. Some reasons for this are: *i)* the probabilistic framework is strongly embedded in the whole risk assessment process (e.g. in the risk practitioner mindset, which is “probabilistic”, in the interpretation and use of results for risk-informed decision-making and, generally, regulatory decisions); *ii)* the perceived lack of empirical foundation of some concepts behind non-probabilistic frameworks [14,15]; *iii)* the difficulty for practitioners to understand (or the failure of the scientific community to communicate) the differences in the mathematical frameworks and their implications in the analysis results [16]. An exhaustive treatment of these issues is beyond the scope of the present paper: refer to [8] and [17] for a comprehensive treatment.

In this context, the present paper takes a practitioner standpoint and compares two frameworks for supporting assessments relevant for risk applications, Bayesian Belief Networks (BBNs) and Fuzzy Expert Systems (FESs), on a specific, very important, issue for Human Reliability Assessment (HRA), namely dependence analysis. Comparison of probabilistic and non-probabilistic frameworks is not a new endeavour and has been addressed in several works, e.g. [8,9,11,17,18]. However, in most cases, the comparison is made at the theoretical level of the mathematical frameworks, which makes its understanding difficult outside, for example Bayesian or Fuzzy Logic experts. The emphasis of the present paper on the application problem is helpful to disseminating, with practical perspective, the understanding of the differences within the risk analysis community.

The application of interest for this paper relates to HRA, the part of PSA analyzing Human Failure Events (HFEs), in terms of the factors contributing to the failure, and quantifying their probabilities, referred to as the Human Error Probabilities (HEPs). Ideally, the models used to derive these probabilities should be built and validated on empirical data – e.g. from operational experience or studies in simulated environments. Examples of this use of data can be found in [19]. However, the challenges in data collection efforts (mostly related to the collection of data related to decision or

diagnosis tasks) are such that in the newest HRA methods (ATHEANA [20], MERMOS [21], CESA [22, 23], MDTA [24]), the source of quantification is mostly expert judgment.

Within HRA, a task that is particularly founded on expert judgment (and highly subjective) is the assessment of the dependence among HFEs. Dependence analysis aims at evaluating the influence of an operator failure in performing one task on the performance of the subsequent task. Intuitively, if two tasks are “closely related” in some fashion and the operators have failed the first task, the failure of the second task may be assessed as more likely than following the success of the first task. From a quantitative point of view, the dependence assessment results in the Conditional Human Error Probability (CHEP), on the basis of the identification of the conditioning influencing factors and the quantification of their impact on the HEP. The dependence assessment may have a significant impact on the risk profile of the system, because the dependent/conditional failure probability may be orders of magnitude larger than the unconditional one.

In current HRA practice, the evaluation of dependence is typically performed through the elicitation of a qualitative dependence level: the Technique for Human Error Rate Prediction (THERP) [25] provides some guidelines for the assessment and also includes formulas for modifying the CHEP according to the dependence level. The direct use of expert judgment to assess the qualitative dependence level has a number of weaknesses, including the traceability and repeatability of the results. In some cases, simplified approaches such as decision trees are used; however the rules represented by these trees do not have a transparent basis [26-27].

A structured alternative is that the evaluation is performed by an analyst, who assesses the conditions of the situation under study through the use of expert systems, i.e., mathematical models which mimic the experts’ evaluation. This way of proceeding should make the evaluation more systematic and repeatable and increase the traceability of the results. Two such expert systems based on Fuzzy Logic (FL) and Bayesian Belief Network (BBN) theory have been previously investigated by the authors, in [28] and [29] respectively. For the purpose of demonstration, the dependence model underlying the two expert systems has been kept simple and it is by no means exhaustive in the inclusion of the factors that may influence the dependence among operator actions in practice.

The goal of this work is to present an in-depth comparison of the two expert system approaches to dependence assessment, based on their application to exemplary case studies. In particular, the focus is on the uncertainty treatment and representation in all the evaluation steps to investigate:

- how the analyst may translate his/her assessment of the situational conditions into suitable input to the expert system for the dependence assessment;
- how uncertainties propagate to the outcome of the dependence assessment;
- how the uncertainty in the dependence outcome is represented.

The scope of the comparison is on the mathematical treatment and representation of uncertainty. Some aspects of a full scope comparison have been left out, most notably, those related to the construction of the two models. In particular, to focus on the mathematical differences, the two models are maintained relatively similar, maintaining a correspondence between the FL linguistic labels and the BBN states and between the fuzzy rules and the Conditional Probability Tables (CPTs) of BBN. A full scope comparison would entail addressing the definition of the linguistic labels and of the BBN states as well as the determination of the Fuzzy rules and BBN CPTs from data, expert elicitation or a combination of the two.

Section 2 briefly outlines the two expert system models previously developed for the dependence assessment; Section 3 describes in details how the uncertainty in the input assessment is handled by the two models and highlights the practical differences. Section 4 focuses on the outcomes of the dependence assessment. Then, in Section 5 the two models are applied to a number of case studies to see how the uncertainties in the input assessment propagate to the output. Conclusions are finally drawn in Section 6.

2 Expert systems for human error dependence assessment

Two HFEs are said to be dependent if the probability of failure of one task changes depending on the success or failure of the other. Since sufficient data are typically not available, the relevant conditional probabilities are qualitatively inferred from the nature of the tasks involved and from human factors considerations. For example, the THERP method [25] proposes guiding principles asserting an increasing level, or degree, of dependence between two tasks if:

- the two tasks are close in time;
- the two tasks share systems and functions;
- the performers of the tasks have similar characteristics.

A discussion of the limitations of current approaches for human error dependence treatment is beyond the scope of the present paper (see [27] for more details). The idea behind the work in [28] and [29] is to build a computable model that represents the relationships among the relevant factors

(for the present application, relevant to the dependence between two events) and between these factors and the HEP. The potential benefits from the use of expert models in HRA (connected with possible increased transparency and repeatability of HRA analysis) have been discussed by the authors of the present paper in [27].

The use of expert systems for HRA dependence assessment here is demonstrated on a working model proposed in [27] and [28] that builds on the qualitative rules suggested in THERP. The model does not include all factors that may impact the dependence level, but is deemed of enough complexity for the purposes of this demonstration.

The inputs to the model are: “Closeness in Time”, “Task Relatedness” and “Similarity of Performers”. “Task Relatedness” is then specified in terms of the “Similarity of Cues” and “Similarity of Goals” (Figure 1). In the subsequent text, the labels of the input and output variables are shortened and referred as “Time”, “Task” and “Performers”. “Task” is then specified in terms of the “Cues” and “Goals” respectively.

Figure 1: Functional relationships among the input factors of the working model of dependence [27]

2.1 The FES for human error dependence assessment

The first expert system considered for the dependence assessment relies on the FL framework [30]. While conventional Boolean logic deals with the concepts which are either true or false, FL models the concepts of partial truth by introducing membership functions (MF). FL has been applied in many decision-making problems and proved to be a promising tool to handle uncertainty, ambiguity, imprecision and vagueness which are inevitable characteristics for many real-life problems [31-33].

The developed FES at the basis of the dependence assessment model is described in detail in [28]. The FES is based on a representation of the expert knowledge in terms of IF-THEN rules, linking the linguistic concepts qualitatively describing the input variables (called antecedents of the rules) to those describing the output variables (consequents of the rules). The rules have to be constructed through a formal expert knowledge elicitation procedure; they serve as a transparent representation of the expert knowledge and support its consistency. To elicit the information needed to build the FES, experts are asked to:

- identify the input antecedents and output consequents;
- partition the ranges of variability of these inputs and outputs into distinct Fuzzy Sets (FSs);
- provide the relationships relating the involved linguistic variables, i.e., the Fuzzy Rule Base (FRB).

The dependence model of Figure 1 explicitly identifies the variables of interest in terms of input and output factors. In the context of the present work, each of the four input factors $x_k, k = 1, \dots, 4$ is qualified in terms of linguistic labels (Table 1).

Table 1: Input factors and associated linguistic labels

Each of the factors takes values on a Universe of Discourse (UOD), here arbitrarily chosen to be $[0,1]$. Correspondingly, to each of the above qualifying linguistic labels is associated a Fuzzy Set (FS), i.e. a support interval on the UOD and a MF that quantifies the degree to which the input factor belongs to a certain label. In this study the shapes and widths of the MFs of a given input factor are chosen so that the resulting partition covers completely the respective UOD. Examples of the partitioning of the UODs of the four input factors by means of trapezoidal MFs are given in Figure 2.

Figure 2: FSs of the input factors “Time” (left) and “Cues”, “Goals” and “Performers” (right)

The UODs of the intermediate “Task” and of the output linguistic variable “dependence level” (“Dependence”) are formed by the labels $y_i = \{\text{ZERO, LOW, MEDIUM, HIGH, COMPLETE}\}$. The output labels correspond to the five levels of dependence introduced in THERP [25]. The labels are associated FSs in a similar way as shown for the input factors. Note that in the dependence model in [27], the analyst does not need to interface with the FES, to ease the use of the model. The judgments are instead given on an anchored scale, without the need for the user to be familiar with the Fuzzy formalism (see for example input assessments in Figure 10).

The Fuzzy rules are built via a systematic expert elicitation process [28]. First of all, a sub-model is constructed to relate the “Task” to the two input factors “Cues” and “Goals”. The expert knowledge concerning the relationships between the two input factors “Cues” and “Goals” and the “Task” output is elicited on pre-specified prototype situations (typically related to extreme values of the factors). The (four) rules elicited from the expert are implemented into fuzzy rules of the form:

1. If Cues is NONE and Goals is NONE then Task is NONE,

2. If Cues is NONE and Goals is COMPLETE then Task is MEDIUM,
3. If Cues is COMPLETE and Goals is NONE then Task is HIGH,
4. If Cues is COMPLETE and Goals is COMPLETE then Task is COMPLETE.

These rules fill the four extreme relations of the FRB underpinning the sub-model of “Task” (Table 2, highlighted cells). To fill in the missing relations, an automatic “interpolation” procedure can be performed for smoothly spreading the consequent labels over the fuzzy rules [28].

Table 2: Complete Table for “Task”, given inputs Cues and Goals

Also for the assessment of the “Dependence”, for each label of the “Task” variable, only the rules associated to the extreme anchor points are elicited and the same interpolation procedure previously adopted is used to fill in the missing rules (Table 3). In total, the process requires direct elicitation of 24 rules - the rest being filled with the interpolation approach.

Table 3: Rules for “Dependence”, given inputs Performers, Time, and “Task = C, H, M, L, N”

Once the FES has been built, it can be used by the analysts to assess the dependence level. The model requires in input the fuzzy sets X'_k , $k = 1, \dots, 4$, representing his/her assessment of the input factors. In this work, a fuzzy inference engine based on the Mamdani inference procedure has been used to mathematically combine the FES rules with the input FS X'_k , $k = 1, \dots, 4$. The mathematical details of the procedure can be found in [30, 34, 48].

2.2 The Bayesian Network for dependence assessment

A BBN is a probabilistic graphical model whose structure consists of nodes linked by directed arcs [35, 36]. Nodes represent random variables and arcs between nodes (linking parent nodes to child nodes) indicate causal or influential relationships between variables. For discrete random variables each node has an associated CPT, containing the Conditional Probabilities Distributions (CPDs) of each possible outcome of the child node conditional to each combination of values of the parent nodes. In recent publications, BBN proved a natural and intuitive model to reproduce the complex factor relationships typical of HRA [37, 38].

With reference to the dependence model of Figure 1, an expert system has been developed in [29], based on a BBN with “ranked nodes” (which represent qualitative variables that are abstractions of

some underlying continuous quantities, typically ranging between 0 and 1 [39]). This type of nodes is used to represent qualitative variables described by a finite set of states or labels. Each of the $m = 4$ input factors of the dependence model (named $x_k, k = 1, \dots, 4$) and the output y are expressed in terms of a BBN discrete node characterized by a set of states $X_k^{v_k}, v_k = 1, 2, \dots, n_k$, for the input factors, and $Y^\eta, \eta = 1, 2, \dots, n_y$, for the output factor. The intermediate factor “Task” is modelled likewise.

To model the relationships among the input and output factors of the dependence model, CPTs have to be defined for the “Task” intermediate node and “Dependence” final node. Generally, the CPTs can be derived using historical data and/or expert knowledge [39]. In the present paper, the approach presented in [39] is adopted: the approach allows deriving the CPTs from the elicitation of a limited number of relationships, expressed by the experts in the same form as the 24 rules elicited in Ch. 2.1. These 24 relationships form the basis expert knowledge, common for both the FES and the BBN, on which the two expert models are built. In the approach in [39], the CPTs of the child node are derived associating a doubly truncated normal distribution (“TNormal”) to the continuous variable underlying the factor labels and discretizing it on the range associated to each label (in the present application the 0 – 1 range is equally split among the labels, i.e. of size 0.2 for a five-label factor, centred in 0.1, 0.3, ..., 0.9). Then, the probability density for the child node is obtained from the function $\text{TNormal}(\mu, \sigma)$, where μ is a weighted function of the input values (on the underlying continuous scale) and σ the standard deviation, representing the degree of uncertainty on the child node value. The child node’s probabilities are derived from a weighted function of the parent node values. Four functions are introduced: Mean Average, Minimum, Maximum, Mix of Minimum and Maximum. The approach is implemented in the Software AgenaRisk (<http://www.agenarisk.com/>), which has been used in the present paper for building the CPTs.

The choice of the appropriate function depends on the effect of the value of the parent nodes on the child node, inferred from the elicited relationships and possibly other qualitative considerations [39]. Note that this choice (along with that of the function parameters, μ, σ) requires a number of subjective assumptions be made, i.e. no hard rules connecting elicited information and these functions exist. The function Mean Average is used for the “Task” node. This function is appropriate when the child node value can be modelled as a weighted function of the parent node values. The four rules relating factors similarity of cues (“Cues”) and similarity of goals (“Goals”) support this assumption. Furthermore, from the four rules, it can be inferred that the influence of “Cues” on the “Task” is larger than that of “Goals” (“Cues” being COMPLETE leads to “Task”

being HIGH, while “Goals” being COMPLETE leads to “Task” being MEDIUM – the other factor being NONE). Regarding the weights, in the present application the weights for “Cues” and “Goals” have been assigned the values of 3 and 2, respectively (these values allow the mode of the CPD in correspondence of the combinations of inputs for the four rules to match the rule output (see Table 4). The weights can be interpreted as degree of correlation between parents and child; in this sense, factor “Cues” has larger correlation with “Task” than “Goals” (this is reasonable from the four elicited relationships). The standard deviation σ of the TNormal distribution represents the degree of uncertainty on the value of the child node (on the underlying continuous scale and therefore, ultimately, on the label discretized scale), given the values of the parent nodes. In the present application, a rather low value is chosen ($\sigma^2=5e-4$, the lowest value usable in AgenaRisk [<http://www.agenarisk.com/>]), finally resulting in the CPDs presented in Table 4. The implications of the choice of a larger value will be discussed later in the paper, see Section 5.3.

The CPTs corresponding to the child node “Dependence” were again derived based on the available elicited relationships (Table 2 and Table 3). The function “minimum” was chosen with weights 5, 1, 1 for nodes “Task”, “Performers” and “Time”, respectively, and again $\sigma^2=5e-4$. The choice of the function “minimum” is intended to represent the effect of the main factor “Task” on “Dependence”, especially in its lowest states (LOW and NONE for which the factor “Task” in practice decouples the actions, i.e. their dependence tends to ZERO, Table 3).

Table 4 gives the CPT derived by applying the algorithm in [39] for the node “Task”, while that for node “Dependence” is reported in Appendix B. Cells shadowed in darker grey indicate the anchor points in Table 2 and Table 3 (i.e. the rules elicited from the expert as presented in Section 2.1); cells shadowed in light gray indicate the consequent of the fuzzy rule corresponding to the same input. Bold font indicates the mode of the CPDs. The general match shown by Table 4 between the CPTs and the Fuzzy rules (i.e. the CPTs are generally peaked on the output of the fuzzy rule) suggests that the two approaches may produce quantitatively comparable results. On the other hand, with regards to the “Dependence” node, Appendix B shows similar results for the higher states of “Task”, while for lower states of “Task” the dependency levels produced by the BBN are consistently higher with respect to those by the FES rules. As shown in Section 5.1, these differences confuse the comparison between BBN and FSS results, shifting the attention away from the focus of the present paper on the mathematical modeling of uncertainty (these differences are not directly due to the uncertainty modeling but by the use of different interpolation approaches to fill in the relationships missing from the expert elicitation). To overcome this difficulty, the comparison cases in Section 5 will be made on a BBN which directly represents the fuzzy rules. For

each combination of the parent node states, (which correspond the antecedents of the fuzzy rules), a conditional probability of 1 is assigned to the corresponding state described in the fuzzy rule consequence. The CPT for node “Task” is reported in Table 5 (the CPT for the output node “Dependence” is built likewise).

Table 4: CPT for node “Task” (TNormal function approach [39])

Table 5: CPT for node “Task” (based on fuzzy rules)

The inputs of the BBN model are probability distributions; more precisely, the analyst must provide a discrete probability distribution $P(X_k^D = X_k^{\nu_k})$ for each k -th input factor, $k = 1, \dots, 4$, where X_k^D is the random variable representing the state of the k -th input factor (D indicates that X_k^D is a discrete variable), $X_k^{\nu_k}, \nu_k = 1, 2, \dots, n_k$ its possible state and n_k the number of possible different states of the k -th input factor.

Providing directly the input factor probability distributions may prove very challenging: the factor state descriptions (in terms of labels such as NONE, LOW, ...) are abstract concepts, so that, in lack of a specific scaling guidance, it is difficult to associate a specific situation to one of them. The anchor-based dependence model interface has been developed in [27] to overcome this difficulty (see Figure 3 and Figure 10). A procedure to convert input judgements on an anchored scale into a format suitable for a BBN has been presented in [29] and is briefly reported here because of its relevance to understand the differences in the FES and BBN results.

As said, each input factor x_k is considered as a continuous variable X_k in an interval arbitrarily chosen as $[0,1]$. On this scale, the anchor points are positioned from 0 (no similarity) to 1 (complete similarity). Figure 3 is an example of this representation for the input factor “Cues”.

Figure 3: Continuous random variable “Cues” with 5 anchor points

Let us, first, consider the case in which the analyst is able to provide an exact value x'_k for the input assessment x_k in the interval $[0,1]$. In this case, it is uncertain which state X_k^D represents the assessment $x_k = x'_k$. This uncertainty is modelled by considering the discrete random variable X_k^D whose possible values are the states $X_k^{\nu_k}$. The probabilities $P(X_k^D = X_k^{\nu_k} | X_k = x_k)$ of being in the

state $X_k^{v_k}$ given an assessment x_k are defined by the parameters $p_{X_k^{v_k}}(x_k)$, with the normalization condition $\sum_{v_k=1}^{n_k} p_{X_k^{v_k}}(x_k) = 1$.

In the case here considered, the values $p_{X_k^{v_k}}(x_k)$ for each value $x_k \in [0,1]$ are obtained by introducing normal distributions $\tilde{p}_{X_k^{v_k}} \sim N(\mu_{X_k^{v_k}}, \sigma_{X_k^{v_k}})$, only considering the values within $[0,1]$, as shown by Figure 4 (left) for a 5-state input factor. Then, the values of $p_{X_k^{v_k}}(x_k)$ are obtained by normalizing the corresponding $\tilde{p}_{X_k^{v_k}}(x_k)$ values (Figure 4, right). The parameters of the normal distributions are set such to replicate the shapes of the memberships functions (cfr. Figure 4 and Figure 2).

Figure 4: Left: Normal distributions $\tilde{p}_{X_k^{v_k}}(x_k)$ for a 5-state variable (only considering values within $[0,1]$); Right: function $p_{X_k^{v_k}}(x_k)$

In the case of a precise single point input assessment x'_k the marginal probability of being in the state $X_k^{v_k}$, i.e., the input of the BBN is calculated as $P(X_k^D = X_k^{v_k}) = p_{X_k^{v_k}}(x'_k)$.

Otherwise, if the input assessment provides in addition an uncertainty interval $[a'_k, b'_k]$, these data are converted into the pdf $f_{X_k}(x_k)$ representing the probability density that the input variable X_k has value x_k . The joint pdf $g_{X_k^D, X_k}(X_k^{v_k}, x_k)$ of being in the state $X_k^{v_k}$ and of having an input assessment x_k is calculated as:

$$g_{X_k^D, X_k}(X_k^{v_k}, x_k) = P(X_k^D = X_k^{v_k} | X_k = x_k) \cdot f_{X_k}(x_k). \quad (1)$$

Then, the marginal probability of being in the state $X_k^{v_k}$ is calculated as:

$$P(X_k^D = X_k^{v_k}) = \int_0^1 g_{X_k^D, X_k}(X_k^{v_k}, x_k) dx_k. \quad (2)$$

The input probabilities propagate to the output discrete probability distribution of the dependence level as:

$$P(Y^D = Y^\eta) = \sum_{X_1^{v_1}, \dots, X_m^{v_m}} \left[P(Y^D = Y^\eta | X_1^D = X_1^{v_1}, \dots, X_m^D = X_m^{v_m}) \cdot \prod_{k=1}^m P(X_k^D = X_k^{v_k}) \right]. \quad (3)$$

The software package AgenaRisk (<http://www.agenarisk.com/>) has been used to perform all the calculation regarding BBNs.

3 Input assessment

In this Section, the representation by the two expert systems of the analyst uncertainty in the input assessment is investigated.

3.1 FES assessment

As previously mentioned, the analyst task in the dependence assessment is to provide to the expert system the evaluation of the conditions of the input factors for the system under analysis. For each factor, the input evaluation is a FS representing the actual condition of the considered variable. In practice, it is advantageous not to require the analyst to be familiar with fuzzy theory and aware of the partitioned UOD underlying the FES. In describing the input factor conditions, the analyst can focus solely on the anchored scale (Figure 3 and Figure 10) to build the input factor FS.

The following guidelines are proposed to convert the analyst information into a FS format suitable for manipulation by the FES, in various instances of assessment of the situation under analysis. If an anchor point applies exactly, the analyst might show his/her confidence in the assessment by providing a precise evaluation of the input factor, without uncertainty; this assessment x'_k is converted into a singleton FS. Figure 5 (top, continuous line) shows an assessment of “Time” exactly halfway between the two anchors “20min” and “5min”. In case of singleton FSs, the BBN input probability distribution is obtained taking the values of $p_{x_k^{v_k}}(x'_k)$.

Figure 5: Conversion of analyst’s assessment for FES and BBN inputs. Top, FES: singleton (solid line), triangular(dotted line) and rectangular(dashed line) FSs; Bottom, BBN: doubly truncated normal pdf(dotted line) and uniform pdf(dashed line)

In other cases, the analyst might be more comfortable in providing not just a point value x'_k to describe the input conditions but also an interval $[a'_k, b'_k]$ reflecting the uncertainty, the variability, or the ambiguity of the situation. Uncertainty relates to the analyst being unsure as to which assessment of the time factor corresponds to the situation under analysis. In other words, the situation is characterized by a specific time separation between two actions, but the analyst is uncertain about the value. Variability relates to the situation being characterized by variable conditions such that the time separation varies within some range (and the analyst is aware of this). Ambiguity relates to the difficulty by an analyst to categorize the situation. For example this may be the case of assessing factor “Cues” (see Figure 3) for temperature parameters measured in two different plant locations. The analysis may be unsure whether these should be considered as two different parameters or a single one. Indeed, generally, uncertainty, variability and ambiguity are not exclusive concepts and are integrated into a single input statement [8].

This situation can be represented by a triangular FS as in Figure 5 (top, dotted line), and a doubly truncated normal pdf $f_{X_k}(x_k)$ could be built (Figure 5, bottom, dotted line).

A similar case is when the uncertainty present in the assessment leads the analyst to provide an interval $[a'_k, b'_k]$ only, without specifying any value of maximum confidence. This evaluation can be translated into a rectangular FS or a uniform pdf, as shown in Figure 5 (dashed line).

3.2 Input modelling comparison

Consider an analyst assessment of the conditions of one of the input factors in the form of the point value $x'_k = 0.2$ with the associated interval of uncertainty $[a'_k, b'_k] = [0.1, 0.3]$. Let us see how the availability of information can change the modelling of the input of both models, considering the three input format presented in Section 3.1.

With reference to the FES framework, Table 6 shows an example of how the three different input fact characterizations would affect the calculation of the degrees of membership to the factor characteristic labels. As expected, by decreasing the quality of input information from a precise single point assessment to an uncertain interval assessment the degrees of membership remarkably differ.

Table 6: Example of input manipulation in the FES and BBN frameworks

Within the BBN uncertainty representation framework, the differences between the modelling through the doubly truncated normal distribution and the uniform distribution can be light. Table 6 shows an example of the input probabilities of being in the state NONE, LOW and MEDIUM for an input assessment $x'_k = 0.2$. As expected a point value assessment gives a more peaked probability distribution while the uncertainty included in a range assessment leads to a distribution spread over more states.

Table 6 shows that the BBN input is less sensitive to less precise information than the FES one, i.e., the differences between the precise single point assessment and the uncertain range assessment are much larger in the FES framework. This is due to the different representation provided by the two models of the input information (input FS or input pdf). The FES computes the degree of activation r_{kj} of the FS X_{kj} of the k -th input variable by the input fact x'_k , whereas the BBN computes the marginal probability of being in the state $X_k^{v_k}$ by performing an integral operation on the joint pdf

$$g_{X_k^D, X_k}(X_k^{v_k}, x_k) \text{ (eq.3).}$$

A further analysis has been performed in order to highlight that these differences do not depend on the different shapes between the fuzzy MFs (trapezoidal) and the elicitation curves (Gaussian) and from the different representation of the input assessment in the case of an interval with a value of maximum confidence (triangular MF and doubly truncated normal pdf), but rather depend from the different mathematical bases of the two input frameworks. To this purpose, consider a BBN model in which the elicitation parameters $p_{x_k^{v_k}}(x_k)$ (Figure 4, right for a 5-states input case) are replaced with trapezoidal ones (Figure 6, left) and the doubly truncated normal pdf of Figure 5 (dotted line) representing a centered range assessment is replaced with a triangular pdf (Figure 6, right).

Figure 6: Trapezoidal elicitation parameters $p_{x_k^{v_k}}(x_k)$ (left) and triangular pdf $f_{x_k}(x_k)$ (right)

The same example of input assessment $x'_k = 0.2$ can be used to test the BBN model with modified input; the results are shown in Table 7.

Table 7: Example of input manipulation in the BBN with modified input

Note that the results and the sensitivity to the amount of information are similar to those found for the original BBN model even if trapezoidal elicitation parameters and triangular pdf are used. In the case study presented in Section 5.1, this modified BBN model is used to clarify the sources of difference between the two modelling frameworks under analysis.

4 Output comparison

The FES and BBN expert systems differ also in the interpretation of the output. The output of the FES is a label membership that represents “how well” the “Dependence” fits the fuzzy description of a particular label; e.g., in the case shown in Figure 7 (left), the output “Dependence” belongs to the label LOW with a degree of membership of 0.89, to ZERO with degree of membership 0.38 and to MEDIUM and HIGH with degrees of membership of 0.25 (the output corresponds to the case study in [28]).

The BBN output is a probability distribution representing “how likely” it is that the variable is in a certain state; e.g., in the case shown in Figure 7 (right), the “Dependence” has a 0.77 probability of being in state LOW, 0.20 of being in state ZERO and 0.03 of being in state MEDIUM.

Figure 7: FES(left) and BBN (right) “Dependence” output

Figure 8: CHEP Probability density functions for each THERP dependence level (unconditional HEP of 0.001, 5tho: $3.3e-4$, 95tho: $3e-3$; lognormal probability assumed [25]) (left); Probability density functions for CHEP weighted on the BBN assessment in Figure 7 (right)

From the FES or BBN outputs on the dependence level, the dependence level (and corresponding HEP) for use in the PSA needs to be determined.

The BBN output probability distribution is easier to interpret and use in the PSA, and allows a clear quantification of the uncertainty included. The THERP handbook [25] suggests values of CHEPs and corresponding confidence bounds (assuming lognormal distributions) for each level of dependence, in correspondence of different values of the unconditional HEP. Figure 8 (left) shows the lognormal probability density functions corresponding to the unconditional probability value of 0.001 (5th percentile: $3.3\text{e-}4$; 95th percentile $3\text{e-}3$). The resulting uncertainty distribution for the CHEP can then be obtained by weighting each probability density function based on the results from the BBN. For example, in Figure 8 (right) the functions are weighted according to the BBN results in Figure 7 (right): 0.2 for the unconditional function, 0.77 for the function corresponding to the LOW dependence function, 0.03 for that corresponding to HIGH dependence. This approach directly produces the figures that need to be included in the PSA: CHEP, 5th %: $6\text{e-}4$, median: $4\text{e-}2$, 95th %: $2\text{e-}1$ (for the BBN assessment in Figure 7 (right)).

In order to evaluate the CHEP within the fuzzy framework, the range of CHEP values is partitioned into the five FSs reported in Figure 9 and corresponding to the five “Dependence” labels. The FSs are centered on the CHEP values suggested by the THERP handbook with support given by their respective lower and upper confidence boundaries (CBs) and have triangular shape in log10 base [28]. The use of the logarithmic scale aids the representation of the different orders of magnitude of CHEP values characterizing the different dependence levels.

Figure 9: FSs for CHEP corresponding to each THERP dependence level (left); FS for CHEP resulting from the FES assessment in Figure 7 (right).

The FS for the CHEP resulting from the FES assessment can then be obtained by associating to each FS of Figure 9 (left), a degree of activation equal to the degree of activation of the correspondent “Dependence” labels (Figure 7 (left)). Thus, the FS reported in Figure 9 (right) is obtained. One can observe that the FES output is less intuitive to interpret than the probability density function obtained by the BBN approach and the degree of membership is not easily convertible into a probability for use in the PSA. However, in the fuzzy framework several techniques have been proposed for the conversion of the final FS into a single value that synthesizes the information encoded in the output FS [40-42]. In particular, a popular defuzzification technique is the middle of maxima (MOM) method which returns the middle value of the core of the FS, being the core the set of values of the output taking the largest degree of membership. In the case of

Figure 9 (right), the most representative CHEP is $5e-2$ which is similar to the median value $4e-2$ obtained in the BBN framework. In order to describe the uncertainty in the CHEP estimation, it is possible to consider the alpha-cut which represent the set of value of CHEP with an associated degree of membership greater or equal to alpha, being alpha a real number between 0 and 1. In this case, the 0.10 alpha cut of the FS representing the CHEP is the interval $[3.7e-4, 9.3e-1]$ which is very different from the $[5^{th} \%, 95^{th} \%) = [6e-4, 2e-1]$ interval which represent the CHEP uncertainty in the BBN approach (this is due to the large membership to the HIGH “Dependence” in the fuzzy assessment of Figure 7, left).

Finally, notice that although the FS defuzzified value and the alpha-cut synthesize the information contained in the output FS, they cannot be directly used in the PSA of the situation under study since their interpretation does not correspond to that of the median and the percentile of the probability density function.

Another option is to consider the possibilistic interpretation of a Fuzzy Set [43-45], which entails defining lower and upper cumulative density functions from which one can obtain a range of values for the probability that the output is in a given interval I of CHEP values.

5 Comparison of FES and BBN models on literature case studies

In this Section, the modelling frameworks for the quantification of human error dependence based on FES and BBN are applied on literature case studies. The examples of Section 5.1 and Section 5.2 are taken from [28]; the example of Section 5.3 is an artefact of extreme situation from the point of view of the uncertainty on the input assessment, for which the two frameworks lead to the largest differences.

5.1 Case study 1: single point estimates input

This case study refers to a set of operator actions required to avoid excessive boron dilution in the reactor cooling system of a nuclear Boiling Water Reactor (BWR) in case of an Anticipated Transient Without Scram (ATWS) scenario [27]. In the considered scenario, the operators have successfully initiated the Standby Liquid Control System (SLCS) to shut the reactor down. To facilitate the reactor shut down, the operators are directed by the procedures to increase the voiding by reducing the level in the reactor to the Top of Active Fuel (TAF). Additionally, they are required to inhibit the actuation of the Automatic Depressurization System (ADS), which is activated by the signal of low water level in the reactor, generated while lowering the reactor water level to TAF. In case of failure to inhibit the ADS, the reactor pressure would be automatically decreased and the low pressure injection systems (Core Spray System, CSS), would be activated. The injected water

could lead to diluting the boron injected by the SLCS and the consequential failure to control reactivity. In case of failure to inhibit ADS actuation, the operators are called to control the level in the reactor using low pressure injection, tripping one of the CSS pumps and controlling the other pump. The signal to activate the ADS is generated about 7 minutes after the event of failure to scram. At that point, the operators have about 15 minutes to take actions to limit the low pressure injection flow.

The pair of operator tasks involved in the dependence assessment are 1) inhibit the ADS and 2) control the reactor vessel level in order to prevent diluting boron concentration after failure to inhibit the ADS. Both actions are directed by the same emergency procedure.

The input judgments (Figure 10) are given as point values on the anchored scales:

- “Time”: as said in the scenario description, the separation in time of the two actions is expected to be around 15 minutes.
- “Cues”: the initial cues for ADS inhibition are related to high reactor power level due to the failure to scram. The cues for control of low pressure injection are related to the reactor vessel level, which has to be manually maintained. The ADS actuation is commanded by the signal of low water level in the reactor; therefore, the level in the reactor is also one of the parameters that the operators have to monitor while inhibiting ADS. The context might then be considered slightly ambiguous. This may be achieved by assigning still a point value assessment but located somewhere in between the anchors “different indicators for different parameters” and “different indicators for the same parameters”.
- “Goals”: the two actions relate to the same function (shut down the reactor by boron control), carried out via different systems. This situation matches the anchor “same function by different systems”.
- “Performers”: the action is carried out by the same operator crew. This matches the anchor “same team”.

Figure 10: Input assessments

Within the FES framework of analysis, the above assessment is represented by singleton FSs (Table 8 and Table 9 report the activation degree of the input labels). The output of the dependence model in form of the possibility values of the different levels of human error dependence is reported in Figure 11 (left). The ambiguity in the input assessment is reflected in the output assignment of a possibility of 0.8 to the LOW level of dependence, 0.2 to the MEDIUM and 0.15 to the ZERO. In the case depicted in Figure 10, four rules are activated:

- IF Time is Neither AND Cues is None AND Goals is High AND Performers is High THEN Dependence is Zero;

- IF Time is Neither AND Cues is Low AND Goals is High AND Performers is High THEN Dependence is Low;
- IF Time is Close AND Cues is None AND Goals is High AND Performers is High THEN Dependence is Low;
- IF Time is Close AND Cues is Low AND Goals is High AND Performers is High THEN Dependence is Medium.

which result in the possibilities of ZERO, LOW and MEDIUM dependence of Figure 11 (left).

In the BBN framework, the input assessment of Figure 10 is represented by single values, leading to the BBN probability distributions of the input factor levels shown in Table 8 and Table 9.

As already mentioned in Section 2.2, two BBNs are considered: one with CPTs derived with the algorithm in [39] (Table 4), and the other with CPTs directly mirroring the fuzzy rules (Table 5). As expected, the former BBN gives rather different results as compared with the FES: Figure 11 shows that the final dependence level assessment by the BBN are shifted by one level compared to the FES. Indeed, as shown in Appendix B, in the CPTs for the higher values of Task, the Dependency tends to be one linguistic term higher than in fuzzy rules. This difference in the relationships underlying the FES and the BBN is determined by the use of two rather different interpolation algorithms for the two modelling frameworks (see Section 2), although they are based on the same set of 24 relationships. Though important, this issue shifts the focus of the present paper, which emphasises the mathematical treatment of uncertainty by the two frameworks (differences in the output results, as in Figure 11, would be due to the differences in the CPTs rather than to differences in the uncertainty representation and quantification). To avoid this, the rest of the comparison cases will be based on the use of a BBN, with CPTs mirroring the fuzzy rules (Table 5).

A qualitative comparison among the outputs obtained by the FES and BBN methods of analysis can be directly performed by looking at the results obtained in terms of the degrees of memberships and the discrete distributions, respectively (Figure 12). As one can see, the two methods give similar results, although the ZERO level has associated a higher probability than the MEDIUM level in the BBN framework whereas the membership of the Zero level is lower than that of the MEDIUM level in the FES framework. The differences among the two model outputs derive mainly from the different shapes between the FES trapezoidal MFs and the BBN elicitation parameters $p_{x_k^{v_k}}(x_k), k = 1, 2, \dots, m$ (Figure 5). In this respect, notice that if the modified BBN model defined in Section 3.2 based on trapezoidal elicitation parameters is used, the obtained output is very similar, except for a normalization factor to the FES output (Figure 12, right, “BBN with modified input”). The differences between the BBN and BBN modified outputs are mainly due to the tails of the

elicitation functions $p_{x_k^{vk}}(x_k)$ that result in different probabilities for some input factor levels. For example, the label “neither” for the input “time” has an input probability of 0.111 (Table 8) in the BBN model, whereas, using trapezoidal elicitation functions in the modified BBN model the input probability of 0.161 is obtained. Thus, it can be concluded that when all the input assessments are represented by single point estimates, no remarkable differences are provided by the inference engines since they behave in a very similar way.

Table 8: FES and BBN input assessments for Time

Table 9: FES and BBN input assessments for Cues, Goals and Performers

Figure 11: Case study 1 output: FES (left); BBN – CPT derived algorithm from [39] (right)

Figure 12: Case study 1 output: FES (left); BBN – CPTs mirroring FES (right).

5.2 Case study 2: uncertain input

This case study shows how uncertainty in the evaluation of the situation under analysis can be represented in the assessment by the analyst and how this reflects in the resulting output dependence. Compared to the case from Section 5.1, uncertainty is considered for the input factor “Time”.

According to the scenario description, the operators have about 15 minutes to take actions to limit the low pressure injection flow. In Case 1, a point value of 15 minutes was used as input. In this case, the linguistic description of the situation “about 15 minutes” is converted in an interval of plausible values between 5 and 20 minutes, with a most plausible value halfway (Figure 13).

Figure 13: Time Input assessment

According to the guidelines given in Section 3.1 for the representation of the analyst uncertainty on the input assessment, a triangular FS is used for the Time assessment (Figure 5 (top, dotted line), input label activation reported in Table 10 and Table 11).

The same rules of the previous Case study 1 are activated although with slightly different degrees of activation. This leads to an output assignment spread on the ZERO, LOW, and MEDIUM levels of dependence with a peak on LOW. In Figure 15 (left), the output of the dependence model in form of the possibility values of the different levels of human error dependence is reported.

In the BBN framework of analysis, the single value input assessments are represented as in the previous Section 5.1 while the input assessment of a range with a most plausible value is converted into a doubly truncated normal pdf.

The calculated joint pdf $g_{X_k^D, X_k}(X_k^{\nu_k}, x_k)$ is shown in Figure 14.

Figure 14: BBN joint pdfs

Then, the BBN probability distributions of the input factor levels are directly calculated and shown in Table 10 and Table 11.

Table 10: FES and BBN input assessments for Time

Table 11: FES and BBN input assessments for Cues, Goals, Performers

The output distribution of the human error dependence level is shown in Figure 15, right. The comparison of the outputs of the two expert system modelling frameworks shows results similar to those obtained in Section 5.1; although the less activated states (ZERO and MEDIUM) have more weight in the FES than in the BBN. The main cause of this difference lies in the different input assessment approach. As described in Section 3, the two expert system models treat differently the input function (input FS or input pdf) to obtain the output: the FES uses the MAX-MIN composition between the fuzzy relations and the input fact while the BBN performs an integration of the joint pdf calculated as the product of the input pdf and the elicitation functions (eqs. 1 and 2).

Figure 15: FES (left) and BBN (right) output of case study 2.

5.3 Case study 3: uncertainties in multiple inputs

Let us now consider a scenario with multiple uncertainties in the input assessment. In particular, input factors “Cues”, “Goals” and “Performers” are assessed to lie between 0.7 and 0.9 while the “Time” is assumed to be 8 hours (Figure 16).

Figure 16: Input assessments

Within the FES framework of analysis, the above assessment is represented by rectangular FSs (Figure 17, input label activation reported in Table 12 and Table 13).

Figure 17: FSs of input facts (solid line)

This activates 8 rules:

- IF Time is Wide AND Cues is High AND Goals is High AND Performers is High THEN Dependence is Low;
- IF Time is Wide AND Cues is High AND Goals is Complete AND Performers is High THEN Dependence is Low;

- IF Time is Wide AND Cues is High AND Goals is High AND Performers is Complete THEN Dependence is Medium;
- IF Time is Wide AND Cues is High AND Goals is Complete AND Performers is Complete THEN Dependence is Medium;
- IF Time is Wide AND Cues is Complete AND Goals is High AND Performers is High THEN Dependence is Medium;
- IF Time is Wide AND Cues is Complete AND Goals is Complete AND Performers is High THEN Dependence is Medium;
- IF Time is Wide AND Cues is Complete AND Goals is High AND Performers is Complete THEN Dependence is High;
- IF Time is Wide AND Cues is Complete AND Goals is Complete AND Performers is Complete THEN Dependence is High.

The FES inference engine provides the output degrees of membership of the human error dependence levels shown in Figure 20, left.

In the BBN framework, the input assessment is converted into the uniform pdfs in Figure 18.

Figure 18: BBN input representation

The obtained BBN probability distributions of the input factor levels are shown in Table 12 and Table 13.

Table 12: FES and BBN input assessments for Time

Table 13: FES and BBN input assessments for Cues, Goals and Performers

The output distribution of the human error dependence level is shown in Figure 20, right.

The comparison of the outputs of the two models shows that a completely uncertain input assessment leads to quite different results: the FES assigns to the three “Dependence” LOW, MEDIUM and HIGH the same strength of 1 while the BBN output distribution is less conservative showing a predominance of LOW and MEDIUM. This affects the analyst final estimate of the dependence level: for example, a conservative evaluation using the FES leads to a confident HIGH “Dependence” while using the BBN model an analyst may estimate a conservative “Dependence” equal to MEDIUM.

The first source of difference is in the different treatment of the input. For example, for the input variable “Cues” the FS “high” and “complete” are equally activated to 1 by the fact FS. This is due to the fact that in the FES it is sufficient to have a single value of the input variable characterized by a degree of membership equal to 1 in both the input fuzzy fact X'_k and in the FS X_{kj} that the degree

of activation of the j -th rule is 1. On the contrary the BBN probability of being in state “high” is 0.734 and in state “complete” is 0.253 (Table 13).

In the BBN framework, the probability of Cues being complete is computed by (eq. 3) and thus depends on the value of the input pdf $f_{X_{cues}}(x_{cues})$ in the entire range of variability of x_{cues} .

The less spread BBN results shows some limitations of the probabilistic representation of epistemic uncertainty under limited knowledge [46]. In this case, for example, the only available information is that the value of the model parameter x_k is located somewhere between $x_{k\min} = 0.7$ and $x_{k\max} = 0.9$. A uniform probability distribution has been used to represent the uncertainty on x_k ; this choice appeals to the Laplace principle of insufficient reason, according to which all that is equally plausible is equally probable, and to the maximum entropy approach [46]. On the other hand, the uniform distribution implies that the degrees of probability of the different values of x_k depend on the equal range $[0.7, 0.9]$ and that a relationship like the following holds:

$$P(x_k \in [0.7, 0.8]) = P(x_k \in [0.8, 0.9])$$

But, if no information is available to characterize the uncertainty under study, then no particular relation between $P(x_k \in [0.7, 0.8])$ and $P(x_k \in [0.8, 0.9])$ should be supported. Indeed, in the most typical BBN formulation, lack of knowledge is represented by specific functional choices of probability distributions, like the uniform. Yet, it should somehow be expressed in terms of the full set of possible probability distributions on $[0.7, 0.9]$ so that the probability of a value $x_k \in [0.7, 0.9]$ is allowed to take any value in $[0, 1]$. This can be done by making the BBN input probabilities dependent on appropriate parameters that allow covering the desired spectrum of probability distributions [36] or by considering the framework of the imprecise probability [47] that allows representing uncertainty by using set of probability distributions.

Annulling the difference in the input treatment of the two expert system models by imposing as input to the BBN the normalized degree of activation of the input FS by the fact X'_k (

Figure 19), the output assessment of Figure 20, right, is obtained. The still large difference is due solely to the inference engines.

Finally, it is worth underscoring the implications of the choice of the parameter σ of the truncated $TNormal(\mu, \sigma)$ distribution capturing the parent-child relationship in the BBN with ranked nodes [39]. σ is the standard deviation of the normal distribution from which the $TNormal(\mu, \sigma)$ is obtained (by truncation within 0, 1 and normalization). The case shown in Figure 11 corresponds to the choice of $\sigma = 5e-4$, the lowest value usable in AgenaRisk (<http://www.agenarisk.com/>). The choice of σ is related to the confidence of the modelled in the uncertainty of the value of the child node as a function of the values of the parent nodes and influences the shape of the BBN CPDs.

Figure 21 shows the BBN output in correspondence of three different levels of uncertainty in the BBN conditional distribution. It is evident that, the larger the parameter σ , the larger the uncertainty, until the (in practice) uniform distribution of the output is obtained.

The same situation of larger uncertainty in the relationship between the parent and child nodes can be represented within the FES framework, by considering multiple fuzzy sets in the consequent part of the rules. Considering, for example, the “Task” assessment, the analyst can decide to use rules such as:

“If Cues is NONE and Goals is NONE then Task is NONE or LOW”

and the complete set of rules reported in Table 14.

Table 14: Complete Table for “Task”, given inputs Cues and Goals in case of large uncertainty on the relationships between “Cues”, “Goals” and “Task”.

Figure 19: Forced BBN input probabilities

Figure 20: FES (left) and BBN (right) output of case study 3

Figure 21: BBN output (input assessment case 3) for different values of the parameter σ of the truncated $TNormal(\mu, \sigma)$ [39]

6 Conclusions

In this work, we have performed an in-depth comparison of two expert systems for the modeling and assessment of the dependence existing in two consecutive HFEs. One expert system is based on a FL framework and the other on a BBN framework, and the focus of the comparison has been on the representation and treatment of the uncertainty associated to the dependence modeling and assessment. To this aim, the following three issues have been analysed:

- 1) how the analyst may translate his/her assessment of the situational conditions into suitable input to the expert system for the dependence assessment;
- 2) how uncertainty propagates to the outcome of the dependence assessment;
- 3) how uncertainty in the dependence outcome is represented.

With respect to 1), we have found that both FL and BBN frameworks can accommodate an anchored scale interface which does not require the analyst to be familiar with the mathematical details. This is an important feature because providing direct input in the proper form for feeding the FES and BBN models can be challenging for the analyst. Furthermore, the comparative analysis of suitably constructed case studies has shown that BBN is less sensitive to less precise information

than FES, i.e. the differences in the uncertainty representation in an assessment based on precise input with one based on an uncertain input, is sensibly smaller in the BBN framework than in the FL framework.

With respect to issue 2) above, by analysis of different case studies we have found that in the case of no uncertainty, i.e., single-value inputs, there is not a significant variability among the outcomes of “Dependence” provided by the two expert systems, BBN-based and FL-based. On the contrary, when the uncertainty in the input is significant, e.g. when a range of values is provided to represent an uncertain input, the FES generates results less certain and more spread on the different levels of dependence than the BBN does. It has been shown that in cases with uncertainties in multiple inputs, these differences may lead to significantly different outputs, with the BBN being less conservative than the FES.

Finally, with respect to issue 3) above, the BBN output probability distributions have been shown to be easier to interpret and use in PSA. In practice, the BBN directly produces the figures that need to be used in PSA, whereas the FES output, which is in the form of a fuzzy set, is less intuitive to interpret and use.

From the considerations drawn from the analysis of issues 1), 2) and 3), we can pragmatically conclude that:

- a BBN approach should be preferred in all the cases characterized by quantifiable uncertainty in the input (i.e. when probability distributions can be assigned to describe the input parameters uncertainty), since it provides a satisfactory representation of the uncertainty and its output is directly interpretable for use within PSA;
- in cases characterized by very limited knowledge, an analyst may feel constrained by the probabilistic framework, which requires assigning probability distributions for describing uncertainty. In these cases, the FES seems to lead to a more transparent representation of the input and output uncertainty.

With respect to this latter conclusion, we underline that a main criticality of BBN emerged in this work relates to the use of uniform probability distributions to represent uncertainty when the information available only indicates that the value of the variable is located in a range of values.

In this conclusion section, we feel that it is also important to remind that the goal of the work was not to arrive at a definite statement of which modelling framework is the best choice for use within PSA. Drawing such conclusion is not possible, in lack of empirical data to validate the results. Rather, the work has illustrated how the two frameworks differently interpret and process the analyst information on the input parameters and how this is transported onto the output, to be used for PSA. For arriving at such use in PSA requires verification of the models (given that validation is not possible), in which the underlying modelling assumptions and output results are scrutinized and

confirmed by experts. Then, the problem shifts to the HRA domain as opposed to the mathematical domain considered here, which is outside the scope of the present work. Indeed, a full scope comparison and verification would entail addressing the definition of the linguistic labels and of the BBN states as well as the determination of the Fuzzy rules and BBN CPTs from data, expert elicitation or a combination of the two.

The emphasis of this paper on the application problem is helpful to disseminating, with practical perspective, the understanding of the differences within the risk analysis community, not limited to Bayesian Belief Networks or Fuzzy Logic specialists. The insights on uncertainty modelling are not specific to dependence and HRA analysis.

Acknowledgements

Part of this work was funded by the Swiss Federal Nuclear Safety Inspectorate (ENSI), under DIS-Vertrag Nr. 82610. The views expressed in this article are solely those of the authors.

The authors would like to thank Michele Conti for running some of the calculations included in this paper.

References

- [1] M.E. Paté-Cornell, “*Uncertainties in risk analysis: Six levels of treatment*”. Reliability Engineering and System Safety, 1996; 54(2-3): 95-111.
- [2] G.E. Apostolakis, “*The concept of probability in safety assessments of technological systems*”. Science, 1990, 250(4986): 1359-1364.
- [3] W. Parry, W. Gareth, P.W. Winter, “*Characterization and evaluation of uncertainty in probabilistic risk analysis*”. Nuclear Safety, 1981; 22(1): 28-42.
- [4] E. Zio, “*Reliability engineering: Old problems and new challenges*”. Reliability Engineering and System Safety, 2009; 94(2): 125-141.
- [5] G.E. Apostolakis, “*Uncertainty in probabilistic safety assessment*”. Nuclear Engineering and Design, 1989; 115(1): 173-179.
- [6] T. Aven, “*Foundations of risk analysis*”. Wiley, NJ, 2003.
- [7] T. Bedford, R. Cooke, “*Probabilistic Risk Analysis. Foundations and Methods*”. Cambridge University Publishing Ltd, Cambridge.

- [8] T. Aven, P. Baraldi, R. Flage, E. Zio, “*Uncertainty in Risk Assessment: The Representation and Treatment of Uncertainties by Probabilistic and Non-probabilistic Methods*”, John Wiley & Sons Inc, 2014.
- [9] T. Aven, “*On the interpretations of alternative uncertainty representations in a reliability and risk analysis context*”. Reliability Engineering and System Safety, 2011, 96: 353-360.
- [10] P. Baraldi, E. Zio, “*A combined Monte Carlo and possibilistic approach to uncertainty propagation in event tree analysis*”. Risk Analysis, 2008; 28(5): 1309-1325.
- [11] D. Dubois, “*Representation, propagation and decision issues in risk analysis under incomplete probabilistic information*”. Risk Analysis, 2010; 30: 361-368.
- [12] R. Flage, P. Baraldi, E. Zio, T. Aven, “*Probabilistic and possibilistic treatment of epistemic uncertainties in fault tree analysis*”. Risk Analysis, 2013; 33(1): 121-133.
- [13] T. Aven, “*Interpretations of alternative uncertainty representations in a reliability and risk analysis context*”. Reliability Engineering and System Safety, 2011, 96(3): 353-360.
- [14] R. Cooke, “*The anatomy of the squizzel. The role of operational definitions in representing uncertainty*”. Reliability Engineering and System Safety, 2004; 85: 313–319.
- [15] W. North, “*Probability theory and consistent reasoning*”. Risk Analysis, 2010, 30 (3): 377-380.
- [16] D.V. Lindley, “*The philosophy of statistics*”. The Statistician, 2000; 49(3): 293–337.
- [17] T. Aven, E. Zio, *Some considerations on the treatment of uncertainties in risk assessment for practical decision making*. Reliability Engineering & System Safety, 2011; 96(1): 64-74.
- [18] B.A. Mrad, V. Delcroix, M.A. Maalej, S. Piechowiak, M. Abid, “*Uncertain evidence in Bayesian networks: Presentation and comparison on a simple example*”. Communications in Computer and Information Science, 2012, 299 CCIS (PART 3), pp. 39-48.
- [19] B. Kirwan, W.H. Gibson, B. Hickling, “*Human error data collection as a precursor to the development of a human reliability assessment capability in air traffic management*”. Reliability Engineering & System Safety, 2008; 93(2): 217-233.
- [20] J. Forester, A. Kolaczowski, S. Cooper, D. Bley, E. Lois. “*ATHEANA User’s Guide*”, NUREG-1880, 2007, U.S. Nuclear Regulatory Commission, Washington, DC.
- [21] P. Le Bot, “*Human Reliability data, human error and accident models – illustration through Three Mile Island accident analysis*”. Reliability Engineering & System Safety, 2004; 83(2): 153–167.

- [22] B. Reer, V.N. Dang, S. Hirschberg, "*The CESA method and its application in a plant-specific pilot study on errors of commission*". Reliability Engineering & System Safety, 2004; 83(2): 187-205.
- [23] L. Podofillini, V.N. Dang, O. Nusbaumer, D. Dres, "*A pilot study for errors of commission for a boiling water reactor using the CESA method*", Reliability Engineering & System Safety, 2012, 109, pp. 86-98.
- [24] J. Kim, W. Jung, Y.S. Son, "*The MDTA-based method for assessing diagnosis failures and their risk impacts in nuclear power plants*", Reliability Engineering & System Safety, 2007; 93(2):337-349.
- [25] A.D. Swain, H.E. Guttman, "*Handbook of human reliability analysis with emphasis on nuclear power plant applications*", NUREG/CR-1278, 1983.
- [26] M. Cepin, "*DEPEND-HRA - A method for consideration of dependency in human reliability analysis*". Reliability Engineering & System Safety, Volume 93, Issue 10, 2008, 1452-1460.
- [27] L. Podofillini, V.N. Dang, E. Zio, P. Baraldi, M. Librizzi, "*Using Expert Models in Human Reliability Analysis – A Dependence Assessment Method Based on Fuzzy Logic*", Risk Analysis, 2010, 8, 1277-1297.
- [28] E. Zio, P. Baraldi, M. Librizzi, L. Podofillini, V.N. Dang, "*A fuzzy set-based approach for modeling dependence among human errors*", Fuzzy Sets and Systems, Vol. 60, No.13, 2009, pp. 1947-1964.
- [29] P. Baraldi, M. Conti, M. Librizzi, E. Zio, L. Podofillini, V.N. Dang, "*A Bayesian Network Model for Dependence Assessment in Human Reliability Analysis*", Proceedings of the European safety and reliability conference 2009, September 7-10, 2009, Prague, Czech Republic, pp. 223-230.
- [30] D. Dubois, H. Prade, "*What are fuzzy rules and how to use them*", Fuzzy Sets and Systems 84 (2) (1996) 169–185.
- [31] L.A. Zadeh, "*The concept of a linguistic variable and its application to approximate reasoning-I*". Information Sciences, 1975; 8(3): pp. 199-249.
- [32] B. Al-Najjar, I. Alsayouf, "*Selecting the most efficient maintenance approach using fuzzy multiple criteria decision making*". International Journal of Production Economics, 2003; 84 (1), pp. 85-100.
- [33] L.T. Tran, C.G. Knight, R.V. O'Neill, E.R. Smith, K.H. Riitters, J. Wickham, "*Fuzzy decision analysis for integrated environmental vulnerability assessment of the Mid-Atlantic region*". Environmental Management, 2002; 29(6):845-859.
- [34] R. Babuska, "*Fuzzy Modeling for Control*", Kluwer Academic Publishers, Massachusetts, 1998.

- [35] J. Pearl, “*Probabilistic reasoning in intelligent systems: networks of plausible inference*”, San Mateo, CA: Morgan Kaufmann Publishers; 1988.
- [36] F. V. Jensen and T.D. Nielsen, “*Bayesian Network and Decision Graphs*”, New York, NY, USA, Springer science, 2007.
- [37] K. Groth, A. Mosleh, “*Deriving causal Bayesian networks from Human Reliability Analysis data: A methodology and example model*”, Proc I. Mech E part O: Journal of Risk and Reliability, 2010, 226(4): 361-379.
- [38] K. Groth, L. Swiler, “*Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H*”, Reliab. Eng. and Sys. Safety, 2013, 115: 33–42.
- [39] Fenton, N. E., Neil, M., and Caballero, J. G., 2007. Using ranked nodes to model qualitative judgments in Bayesian networks. IEEE Transactions on Knowledge and Data Engineering, Vol. 19, No. 10, pp. 1420–1432.
- [40] G. Shafer, “*A Mathematical Theory of Evidence*”, Princeton, NJ: Princeton University Press, 1976.
- [41] D. P. Filev, “Fuzzy models induced by alternative defuzzification methods”, Proceedings of the fifth IEEE international conference on fuzzy systems, 1996.
- [42] C. C. Lee, “*Fuzzy logic in control systems: Fuzzy logic controller. Part II*”, IEEE Transaction on Systems, Man and Cybernetics, 20(2), 1990.
- [43] W. Leekwijck and E. E. Kerre, “*Defuzzification: Criteria and classification*”, Fuzzy Set and Systems, 108, 159–178, 1999.
- [44] D. Dubois and H. Prade, “*Possibility Theory: An Approach to Computerized Processing of Uncertainty*”, New York, Plenum Press, 1988.
- [45] G.J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Upper Saddle River, NJ, Prentice-Hall, 1995.
- [46] H. Gzyl, “*The Method of Maximum Entropy*”, Ser. Adv. Math. Appl. Sci., Vol. 29, 1995.
- [47] P. Walley, “*Statistical Reasoning with Imprecise Probabilities*”. Chapman & Hall, London, 1991.
- [48] C. Bilgi and T. Taner, “Measurement of membership functions: theoretical and empirical work”, Fundamentals of fuzzy sets, 195--227, Springer, 2000.

Appendix A: CPT elicitation algorithm using ranked nodes

This appendix presents some details of the approach used for derivation of the BBN relationships (i.e. the CPTs). For a more comprehensive treatment, see [39].

The approach is applicable for BBNs with “ranked nodes”, which represent qualitative variables that are abstractions of some underlying continuous quantities, typically ranging between 0 and 1). The key concept for the approach is that the child node’s probabilities (each CPD making the CPTs) are derived from a weighted function of the parent node values (on the underlying continuous scale).

The CPDs of the child node are derived associating a doubly truncated normal distribution (“TNormal”) to the continuous variable underlying the factor labels and discretizing it on the range associated to each label (in the present application the 0 – 1 range is equally split among the labels, i.e. of size 0.2 for a node with five levels, centred in 0.1, 0.3, ..., 0.9).

Then, the probability density for each child is the function $\text{TNormal}(\mu, \sigma)$, where μ is a weighted function of the input values (on the underlying continuous scale) and σ the standard deviation, representing the degree of uncertainty on the child node value. Four weighted functions are introduced: Mean Average (W_{mean}), Minimum (W_{min}), Maximum (W_{max}), Mix of Minimum and Maximum (W_{minmax}).

The decision of which function to use depends on the effect of the parent nodes values on the child node value. As presented in [39], this can be inferred from limited information, e.g. specific evaluations in correspondence of combinations of input values (typically, the cases where the nodes have their extreme states). For example, in the case of one child node Y with two parents X_1 and X_2 , if the following statements are elicited from experts [39]:

- when X_1 and X_2 parent nodes are both ‘very high’ the distribution of Y child node is heavily skewed toward ‘very high’,
- when X_1 and X_2 parent nodes are both ‘very low’ the distribution of Y child node is heavily skewed toward ‘very low’,
- when X_1 is ‘very low’ and X_2 is ‘very high’ the distribution of Y is centered below ‘medium’,
- when X_1 is ‘very high’ and X_2 is ‘very low’ the distribution of Y is centered above ‘medium’,

then it is appropriate to use the weighted average function (with possibly different importance weights for the two parents). A simple weighted sum model is used to measure the contribution of each parent node to explaining the child node as a ‘credibility weight’. The higher the credibility value, the higher the correlation between the parent node and the child node. The weights are derived from judgment. Mathematically, for child node Y , having $X = \{X_1, X_2, \dots, X_n\}$ causal ranked nodes as parents and each X_i parent node having w_i contribution weight, the TNormal distribution with weighted mean average function will have the following form:

$$p(Y|X) = TNormal \left[\frac{\sum_{w_i=1}^n w_i \cdot X_i}{\sum_{w_i=1}^n w_i}, \sigma, 0, 1 \right]$$

As mentioned before, also other weighted rank node functions can be used to derive the probability values in CPT. The following observation, for example, will lead to the use of weighted minimum function :

- When X_1 and X_2 parent nodes are both ‘very high’ the distribution of Y child node is heavily skewed toward ‘very high’.
- When X_1 and X_2 parent nodes are both ‘very low’ the distribution of Y child node is heavily skewed toward ‘very low’.
- When X_1 is ‘very low’ and X_2 is ‘very high’ the distribution of Y is centred toward ‘very low’.
- When X_1 is ‘very high’ and X_2 is ‘very low’ the distribution of Y is centred toward ‘low’.

The corresponding function will have the following form:

$$p(Y|X) = TNormal [Wmin, \sigma, 0, 1]$$

With Wmin:

$$Wmin = \min_{i=1, \dots, n} \left[\frac{w_i \cdot X_i + \sum_{j \neq i}^n X_j}{w_i + (n - 1)} \right]$$

If all the weights are large, then Wmin is close to the minimum value of the inputs, and if all the weights are 1, then Wmin is the average of the parent nodes (Wmean). Mixing the influence of the weights gives result between MIN and AVERAGE.

Function Wmax operates analogously.

The approach is implemented in the Software AgenaRisk (<http://www.agenarisk.com/>).

Appendix B: CPT for Dependency node

Table 15: CPT for Dependency node with TNormal function approach [39]